DOCUMENT RESUME

ED 161 931 TM 007 958

AUTHOR Donlon, Thomas F.; Widiger, Thomas A.
TITLE Pictorial Sequences As the Basis for Tests of Reasoning.
PUB DATE [Mar 78]
NOTE 37p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (Toronto, Ontario, Canada, March, 1978)

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS Cartoons; *Cognitive Tests; Culture Free Tests; Factor Structure; *Historical Reviews; Intelligence Tests; *Nonverbal Tests; Pictorial Stimuli; Test Bias; *Testing; Test Validity; Thought Processes; *Visual Measures
IDENTIFIERS *Picture Sequencing Format

ABSTRACT

The historical development and recent applications are described of the picture sequencing format (PSF), a testing device which requires the subject to designate a complete and correct order for a set of related pictures. The earliest use was by Decroly in 1914 and perhaps the best known use is in the Wechsler intelligence tests. Special characteristics and qualities of the format are discussed. The author discusses four major areas within which the PSF may prove useful: (1) for the purpose of diagnostic or differential assessment; that is, broadly testing different areas of specific abilities in order to better grasp the pattern or profile of an individual's attainments; (2) for the testing of certain minority groups, for example the deaf and ethnic minorities; (3) as a predictive measure of certain practical, real life criteria, such as job success; and (4) as an intuitively appealing and entertaining measurement device useful in combatting low motivation, test anxiety, and negativism. (Author/CTM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Thomas Donlon

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

ED161931

Pictorial Sequences

As The Basis For Tests of Reasoning

Thomas F. Donlon
Educational Testing Service

Thomas A. Widiger
Miami University
Oxford, Ohio

National Council on Measurement in Education

Toronto, Canada

Spring, 1978

TM007 958

Pictorial Sequences

As The Basis For Tests of Reasoning

Thomas F. Donlon
Educational Testing Service

Thomas A. Widiger
Miami University

The Picture Sequencing Format (PSF) is a testing device which requires the subject to designate a complete and correct order for a set of pictures. The criteria for completeness and correctness are that the designated series "tells a story which makes sense." The best-known applications of the PSF are the Wechsler Tests, the WAIS and WISC-R, which include Picture Arrangement subtests requiring a physical rearrangement of a set of cards. There have been several variations on the basic theme throughout its history in intelligence testing. In group administrations, for example, the set may be presented in scrambled order and subjects are required to write the correct number beneath each one. Or subjects may be asked to designate a replacement for a missing picture, drawing upon a set of possibilities. This paper refers to all those variations as Picture Sequencing Format.

Historical Development

The Picture Sequencing Format has had a long but inconsistent history. Its earliest usage as a testing device can be traced to the work of O. Decroly (1914), Vermeylen (no date available), and the experience of "Miss Bowler, Whipple, and others' (as cited by Yerkes, 1923).

Based on these sources, as well as their own inventiveness, the psychologists and consultants to the U. S. Army in World War I developed picture arrangement subtests for the Beta and Performance (but not Alpha) forms of the well-known Army mental tests.

The Beta was conceived as a group test of intelligence, given without verbal instructions, and devised for poor readers and illiterates. In this context it was thought that a PSF format would be a good substitute for the typical verbal item (Yoakum and Yerkes, 1920). However, due to its difficulty in administration through pantomime directions, as well as its similarity to, and overlap with, the simple picture completion subtest, it was removed from the final form (Yerkes, 1921, pp. 372 and 374).

It appears that these were the first major published forms for a PSF and the content as well as style of the PSF's which followed greatly borrowed from them. It becomes evident as the history is traced, however, that it was difficult for experimenters to be satisfied with their test. There is a continual dissatisfaction with the pictures finally chosen as well as with the structural format. This dissatisfaction is frequently coupled with a vague realization of a further potential if more time and money were available for the creation and construction of items. For example, Yerkes wrote: "The pictures selected...are far from satisfactory, but they were the best immediately available. Scoring offered some difficulty and the method adopted is not entirely satisfactory" (Yerkes, 1921, p. 401). Such comments and criticisms of the PSF are representative of the feelings generated by a number of authors who attempted to use the PSF.

In spite of the problems, however, Otis (1929), included the format in the group intelligence test he developed after the war. He considered it appropriate, however, only within the Primary Exam for younger subjects,

not the Advanced Exam.

Nor, in fact, did the Army entirely give up, for they included a PSF within their Performance test. The Performance test was individually administered and given to those for whom the Beta results indicated a need for a more detailed, and even less verbal, examination. The items chosen for the Performance-PSF series were not, however, the same as those in the preliminary Beta. In contrast to the more serious, practical stories in the pre-Beta (e.g., man going to jail for crime, house catches on fire, boy hitting baseball, etc.) the series was based on the Foxy-Grandpa cartoons used earlier by Miss Bowler and Dr. Whipple (Yerkes, 1929). These concerned the mischievous and humorous deeds of two young boys and their clever, "foxy" grandfather.

Yerkes' dissatisfaction, quoted above, refers to the Foxy-Grandpa series and its scoring system, and this test is often cited as a disappointing failure. But the early criticism of the Foxy-Grandpa series (unsuitable for adults, too subject to large chance variation scores, and not very discriminative among subjects) seems more a reflection of the type and limited number of series chosen than any general characteristic of the picture arrangement format. The PSF does have certain difficulties, most of which were noted in the Army test development, but these difficulties do not seem inherently insurmountable. There was a general lack of certainty and confidence in the Army with the ultimate decisions as to the choice of content for the series (e.g., moral story or humorous cartoon) and the determination of a scoring system (e.g., "will there be credit given for alternate arrangements and time bonuses?"). Indeed, the variations on these themes result in a multitude of possibilities; and the dissatisfaction lay

more in not knowing how to decide than in any deficiency in performance. In fact, the PSF correlated with Binet Mental age .72 and with the total Performance score .86, both values ranking the third highest among the total of ten performance subtests.

Cornell and Coxe (1934) recognized the potential of a PSF and included a series within their test, the Performance Ability Scale. It again obtained very respectable validity coefficients. Correlation with total performance equaled .84, while correlations with other subtests ranged from .58 to .72. Even when age was held constant, the range of coeeficients was .24 to .49.

Brown, Stein, and Rohrer (1936) developed a picture arrangement series for the Chicago Non-verbal Examination, a test developed for the hard-of-hearing and the deaf. These subjects seemed to overcome the "handicaps" of instruction by pantomime which had proved to be such a drawback in the Army Beta test.

In spite of the attention given to group strategies in the '20s and '30s, interest in them seems to have died. Only within the individually administered Wechsler-Bellevue, WAIS, and WISC-R is a picture arrangement format still used today. It no longer appears within any standard group administered intelligence test. Wechsler continued the tradition of expressing dissatisfaction with the final selection: "...in spite of the considerable labor spent before definitive choices were made, the final selection leaves much to be desired..." (p. 208), (Matarazzo, 1972) but final correlations with full scale are still respectable, ranging from .62 to .77. Reliability and inter-subtest correlations for the WAIS PA tend to be erratic, the former being among the lowest for the WAIS subtests (ranging from .60 to .74 in Wechsler's

data, from .18 to .86 in other studies cited within the manual). This anomaly, however, may be a function of using so few items with a test that involves so many vicissitudes. An expansion and deliberate variation of the qualities of the PSF within a large, group administered test should avoid this instability.

It is interesting to note Wechsler's choice of content for the Picture Arrangement (PA) items, in that many were adpated from the Army Pre-Beta. In fact, the close similarity of the pre-Beta, Otis, Chicago Non-verbal, Cornell-Coxe, and W-B/WAIS items, at times approaches replication. This interesting phenomenon can be taken as a reliance on the "tried and tested" but it is also an indication of the lack of an attempt to experiment with and broaden the format. Wechsler <u>did</u> develop new items for the WISC-R, replacing the Foxy-Grandpa by the Little King, Louie, and other cartoons, and updating the characters and illustrative style, but the fact that the "set of pictures included in [the] battery represents the final choice from among more than twice that number" (Matarazzo, 1972, p. 208) does not really seem to be much of a broad consideration. It probably represents less than a total item pool of twenty-five items.

In reviewing the characteristics of Wechsler's tests, Matarazza again repeated the consistent dissatisfaction with the development effort which was coupled with the following belief in its potentiality:

> In spite of certain definite limitations, the Picture Arrangement test has some very worthwhile merits. In the first place, it is the type of test which effectively measures a subject's ability to comprehend and size up a total situation. The subject must understand the whole, must get the "idea" of the story...Secondly, the subject matter of the test nearly always involves some human or practical situation. The understanding of these situations more nearly corresponds to what other writers have referred to as "social intelligence." (p. 208-209)

The PSF is a complex stimulus involving numerous dimensions. It includes not only an identification and analysis of the stimuli within each panel, but also an integration and global, contextual appraisal of their interrelationship. As will be indicated below, knowledge concerning scientific, moral, and social relationships, as well as short-term memory, analytic, and synthesizing abilities, are all appraised within a deceptively simple and entertaining format. An expansion and deliberate variation of the qualities of the PSF within a large, group administered test should remove many of the disadvantages which have been cited, since these disadvantages seem to arise largely due to the usage of limited numbers of poorly chosen items.

Validity Studies

Exactly what a PSF measures is rather difficult to state. This is largely due to the fact it is a complex, multi-dimensional stimulus which has been evaluated primarily with short tests of poorly (or quickly) chosen items. For example, most of the research has been concerned with the Picture Arrangement Subtest (PAS) of the Wechsler scales. This experience is extensive in terms of the number of individuals considered in samples, but is extremely limited when viewed as a sampling of the content and format domains of this problem type.

On the basis of the classic factor analytic work with the Wechsler scales by Cohen (1959), Sattler (1974) states that "The Picture Arrangement subtest is the best single measure of 'g' among the Performance subtests... In combination with the Block Design subtest, it can be used to obtain a nonverbal estimate of 'g'" (o. 183). The best overall measures of "g" are the Vocabulary and Information (i.e., verbal) subtests, but it is

interesting to see a PSF test emerge superior to other Performance measures.

PA loads on the "Freedom from Distractibility" and "Quasi-Specific" factors in Cohen's analysis. The latter factor "defies interpretation" (Cohen, 1959), so that it does little to describe PA. In Cohen's own rearrangement of the subtests to form his new factor measures, PAS is not included except as one among many (i.e., all of them) in a measure of "g".

The results of factor analysis of Wechsler's PA tests are far from unequivocal. Kaufman (1975), for example, in an extensive factor analysis of the WISC-R, found that the picture arrangement subtest did not load highly on any of the Cohen factors. In fact, contrary to Cohen, he placed it within the Perceptual Organization factor rather than Freedom from Distractibility. As Klinger and Saunders (1975) point out, the use of Wechsler's subtest scores in factor analysis assumes that the items within each subtest are factorially congruent. With such a complex stimulus as a PSF this may be asking a good deal. For instance, the "FLIRT" item in the WAIS is often cited as a unique item distinct from the rest of its subtest (Breiger, 1956; Klinger and Saunders, 1975; Schill, 1966). At times this FLIRT example demonstrates relationships with certain variables that are in opposite direction with the rest of the subtest's items (Blatt and Quinlan, 1967).

To test the homogeneity assumption, Klinger and Saunders (1975) factor analyzed the items of nine subtests of the WAIS. Unlike most of the other subtests, PA was split into two distinct factors: "planning ability or anticipation" and "social awareness." This major substructure within the subtest might go far to explain its variable loadings in factor analytic studies.

Klinger and Saunders' two factors are congruent with the major clinical speculations concerning the PSF. In regard to planning ability and anticipation, Rapaport, Gil, and Schafer (1968) found a marked impairment in the PAS with schizophrenics, impulse disorders, and depressives, the three clinical groups noted for their impairment in regard to future extension. Considering this fact, they concluded that the PA is a measure of the ability to extend one's thinking into the future. Joslin (1966), using Porteus Mazes, Dickstein and Blatt (1966; 1967), using accepted projective measures of future extension, confirmed the subtest's measure of anticipation.

The "social awareness" factor, on the other hand, is the area in which there has been greatest interest and research. Wechsler (1958), did not believe that a "social intelligence" factor existed distinct from "general intelligence applied to social situations" (p. 75). Work on the topic has been somewhat varied and erratic. Winne and Gittinger (1973), for example, developed a Personality Assessment System which involves an analysis of the WAIS in terms of personality factors. One factor, called Role Adaptivity, is an index of a person's skill in meeting role demands, and its major measure is the PAS. A high scorer is assumed to be charming, socially adaptive, makes good impressions, etc. However, Turner and Horn (1976a; 1976b; 1976c) using various personality measures, Krippner (1964) using the Vineland Social Maturity Scale, and Joslin (1966), could not validate this interpretation.

Studies which have analyzed PA scores for actual subgroups of the population deemed to vary on a dimension of social awareness, have also found mixed results. McDonald and Crookes (1967) found lower scores on the PAS (relative to other subtests) for British psychiatric patients.

They attributed this to the group's lower degree of social alertness. Solway, Fruge, Hays, Cody, and Gryll (1976), Solway, Cook, and Hays (1976), and Solway, Hays, Roberts, and Cody (1975), have found both supporting and discounting results when comparing the scores of delinquent psychopathic and nondelinquent adolescents. The distinction seemed to be that the WISC-R PAS items were differentiating while WISC items were not. This would seem to be quite a subtle distinction for such differing social groups.

Finally, Schill (1966) found that the PAS was related to the MMPI scale of social introversion (although not replicated by Turner and Horn, 1976b). He also found that it predicted verbal conditionability (i.e., responsiveness to social stimuli) and the tendency to participate in extracurricular social activities (Schill, Kahn, and Muehleman, 1968a; 1968b).

Probably the most extensive application of a PSF to the social realm has been undertaken by Guilford, O'Sullivan, and Hoepfner. In an attempt to validate Guilford's (1959) Structure of Intellect model, they have created a number of tests which attempt to measure the various social abilities that are associated with the crossing of the five major operations and six products in the model. Other authors have developed tests to measure these factors within the figural, symbolic, and semantic areas, but Guilford also proposed a social or behavioral realm distinct from the other three.

To measure the social intelligence factors, Guilford's project developed a number of tests using cartoons and photographs of expressions, social interactions, and gestures. Examples of Guilford Tests would be: 1) Missing Cartoons: In this test three panels are presented in order,

and a fourth is needed to complete the sequence. The subject must choose which among four possibilities would belong in the empty fourth slot. It is considered a measure of "cognition of behavioral systems;" 2) Cartoon Predictions: In this test a single panel is presented, and the subject must choose from among three others the one which presents the next most likely occurrence. It is seen as a measure of the "cognition of behavioral transformations;" 3) Cartoon Exchange: In this test, the subject must choose the one of four alternative cartoons which, when substituted for a designated panel in an intact strip, changes the meaning of the whole story. This is considered to measure the "cognition of behavioral transformations." There are certainly many others but the above sample is indicative of the potential variations in the structure of a PSF, and the varied content possibilities within and across panels.

It is interesting to note that the PAS (à la Wechsler) is considered by O'Sullivan and Guilford (1975) and Meeker (1969) as being within the semantic rather than the social realm. In their view, it is not simply the use of cartoons showing people which makes a social test, for if the situations are verbally defined, semantic abilities will provide the dominant variance in solving the task. O'Sullivan and Guilford (1975) conducted an extensive factor analysis of their behavioral tests, including a number of marker tests of other factors. Among the latter was Adkins and Lyerly's (1951) Picture Arrangement test (using Louie cartoons in the manner of Wechsler). Adkin's and Lyerly's test loaded on a "convergent production of semantic systems" factor, rather than any of the behavioral. This result again notes the complexity and diversity of the PSF. Guilford's tests are very well chosen, each being

specific to the factor it wishes to measure.

In O'Sullivan and Guilford's factor analysis the validity of the social dimension was only moderately supported. (Hoepfner, 1973; Hoepfner and O'Sullivan, 1968; Shanley, Walker, and Foley, 1971; Tenopyr, 1967). The work relies solely on intertest correlations. There have been no attempts to relate the so-called social intelligence tests to relevant external criteria. Some sort of construct validation, using convergent and discriminative measures, based on academic versus social criteria, would be a useful extension.

The factor analytic studies of Wechsler's and Guilford's tests are the primary validity work to date. Examples of other efforts toward a broader understanding of the PSF are few. One interesting development is the application of a PSF (The Wechsler PAS from the WISC) to Piagetian studies. In two studies which compared the subtest's ability to differentiate groups of children on Piagetian tasks (Orpet and Meyers, 1970) and on conceptual analytic ability (Kagan, Rosman, Day, Albert, and Phillips, 1964), only the PAS (versus other Wechslerian subtests) was able to differentiate between the groups. Such analyses are relative to the subtests considered, the separate reliabilities of the scales, and the tasks considered, but they do indicate a worth and potentiality for PSF perhaps not seriously considered before.

The Characteristic Quality of PSF

The true importance of the Guilford studies, however, lies perhaps in their ability to develop useful PSF tests without the seemingly insurmountable administrative headaches of the earlier developers. Their willingness to break tradition and expand on the strict, PA format

(while retaining the social settings) has probably been their most productive development. Certainly Missing Cartoons and Cartoon Predictions present an administrative procedure far easier than the cumbersome rearrangement of a series of independent cards into alternative orders. What needs to be accomplished, however, is a more systematic variation of the content and complexity of the social scenes themselves.

In other words, it seems to be the use of pictorial social scenes across time, within a meaningful context, which gives the PSF its distinctive flavor and greatest potential. Certainly there is no lack of pictorial tests of intelligence. But such tests tend to use pictorial material either as abstract symbols or as single concrete items taken out of a meaningful, realistic context. Examples abound: Dunn's (1965) Peabody Picture Vocabulary Test (PPVT) and Ammon's Full-Range Picture Vocabualry Test are well-known. These tests are heavily semantic. The PPVT, for example, tends to correlate more with the verbal scale of the WISC than the performance (Sattler, 1974). French's Pictorial Test of Intelligence (1964) contains a picture-vocabulary subtest as well as a number of other tests which are essentially translations of WISC verbal skills into pictorial form. The SRA Pictorial Reasoning Test (McMurray & Arnold, 1973) uses the same format as the Philippine Nonverbal Intelligence Test (Guthrie, Tayag, and Jacobs, 1969) and the Columbia Mental Maturity Scale (Burgemeister, Blum, and Lorge, 1954), in which a set of pictures of single items are presented, the subject's task being to identify which one does not belong in the whole set. The fact that it makes no difference whether the items are of abstract or concrete figures rules out any potential contribution through the use of contextually meaningful social stimuli. Rulon's (1966) Semantic Test of Intelligence

and Gilliland's Learning Potential Examination (1966) both require the arrangement of symbols to tell a story, but again they involve contextually-free, abstract symbols, depending largely on a verbal and semantic understanding.

The list can go on, but it is evident that most present pictorial tests tend to involve format (analogies, similarities, identification of missing parts, etc.) and content which make them more similar to verbal scales and abstract tests (e.g., Raven's matrices tests) than to tests involving practical, socially meaningful stimuli. Such tests do share in the measurement of "g" as well as certain other specific abilities, but they are in no sense expansions and developments of the PSF.

Potentialities of a PSF

In the authors' view, there are four major areas within which the PSF may prove useful: 1) For the purpose of diagnostic or differential assessment; that is, testing broadly different areas of specific abilities in order to better grasp the pattern or profile of an individual's attainments; 2) for the testing of certain minority groups, as for example, the deaf and ethnic minorities; 3) as a predictive measure of certain practical, real life criteria, such as job success; and 4) as an intuitively appealing and entertaining measurement device useful in combatting low motivation, test anxiety, and negativism. Each of these facets will be considered in greater detail in the following discussion.

1) Diagnostic Measurement

The potential for diversity in measurement using variations on the PSF has been noted above. In terms of structure, the Guilford group has made excellent advances. Their variations allow for distinctive assessments

of the ability to anticipate consequences, to redefine behavioral relationships, to organize complex, interrelated information, and so forth. These mental processes involve not only an anlysis of the details within each picture panel, but an integration or synthesis of these separate facts into a meaningful relationship across time. In terms of content, four broad taxonomic categories can be identified in the work on PSF's to date: moral-ethical themes, time-relationship themes, social and interpersonal sensitivity, and sense of humor.

In the realm of "moral maturity" or "knowledge of moral and ethical norms" the use of serial, pictorial images might make the presentation of problem situations far simpler and more direct than the use of a verbal medium. Some of the earlier tests mentioned above made an initial start in this area. The Chicago Nonverbal, Army pre-Beta, and WAIS all contained a four-panel sequence concerning a man who commits a crime, gets arrested, goes to court, and is put in jail. (The W-B and WISC-R include a similar series on an arrested burglar.) A series about curtains catching on fire which leads to a house burning down is found in the pre-Beta and Otis tests. These series have no human agents. But in the "moral" WISC-R version, the series is expanded to introduce a boy playing with matches. His regret is clear in the final panel. Of course, the mischievous pranks of all children (e.g., in the Foxy-Grandpa series) meet with retribution. These existing examples would not be adequate for a complete scale (or perhaps even part of one), but they do illustrate where expansion may be useful. In our view, one may use PSF techniques to approach moral development much as Guilford and his colleagues approached social intelligence. What is needed is a broad, group-administered PSF test with a large, fully saturated moral-ethical subtest.

Another diagnostic dimension could be temporal relationships across a broad array of physical and social events. This has been the dominant theme within PSF testing (e.g., Blatt and Allison, 1968). The 1919 Otis test, which seemed to dismiss moral content, concentrated on pure temporal relationships to the point of including items which consisted only of single pictures which were ordered on the basis of quite sweeping time-change considerations (e.g., the face of a female at various ages in her life, different modes of transportation at various period in history, etc.). The pre-Beta did not include such "sweeping" items, no do any of the Wechsler tests. The absence of any action connecting or continuing across the various panels, as well as the lack of any developed, meaningful context within any panel, would probably make such items more related to the present-day "semantic" pictorial tests (e.g., SRA Pictorial Reasoning, Raven's matrices, etc.) than to the more behavioral and comical stimuli of typical PSF items. But the question is an empirical one, and there is no present evidence (other than the Guilford work) which bears on it. The fact that the item-type has since appeared only in the Chicago Nonverbal (with essential replication), however, indicates its potential theoretical distinction from other types of PSF tests.

The type of "time-relationship" item which seems to be most widely used is that which involves typical everyday progressions (e.g., get up, eat breakfast, go to work). Sometimes the answer is to be found through the knowledge of the progression within the real world (e.g., one knows that corn is sold after it is harvested), while in other cases the order is given more by specific pictorial clues which are usually details within a frame (e.g., changes in the hands of a clock). These two alternate strategies could be quite distinct, for the former appears

to be based on an obtained knowledge of practical, real world events, while the latter is highly dependent upon a visual, analytic scanning strategy.

Other formats for time-relationships not found within the present tests can be developed. For example, the "GURU" cartoon (see appendix) by Kovarsky contains an interesting arrangement in which two panels are exactly alike. Since they depict the silent musings of a mountain guru, one infers the passage of a great amount of time coupled with deep thought. This interaction between the format and content presents further possibilities.

A third diagnostic facet lies in the PSF's potential for measuring "social intelligence." An important facet in this assessment is that similar behavioral gestures and expressions can mean different things within different contexts (Hoepfner, 1973) and it is precisely the usage of a developed, sometimes elaborate, social context that distinguishes the PSF from other tests. An attempt to measure social sensitivity, or the ability to interpret facial and behavioral expressions (i.e., empathy), without considering that in real life these gestures are meaningfully embedded within some context, is likely to be fairly inadequate in predicting outside criteria. Heussestamm and Hoepfner's Inter-Person Perception Test (1969) which was an attempt to develop culture-specific abilities for the interpretation of facial expressions, is a good example of such difficulties. That the PSF possesses a greater potential is clear, for it is the PSF's ability to depict social contexts within and across frames that makes it a truly meaningful and interesting measurement of knowledge or intelligence within the social realm.

A "final" diagnostic area concerns humor, especially social (human or animal) comedy. Sensitivity to humorous situations and acts, the ability

to catch a joke or identify the pun, is an interesting construct. Clearly, humor has been an important aspect in PSF testing. The Army Performance scale consisted entirely of the Foxy-Grandpa cartoons (Wechsler's Little King and Louie series are modern, but subdued examples), and although it would seem to be the time relationship across the panels which was the primary basis for their measurement properties, the humorous nature of the items must have contributed some unique variance. In our informal testing of PSFs, using comic strips from various magazines (e.g., Kovarsky's World), subjects tended to solve many items by first identifying a key panel then building around it. As in Guilford's Cartoon Exchange subtest, one key panel determined the theme or meaning of the whole series. In a major proportion of cases this key panel was the "pun" scene. The trick to solving the arrangement was to first "get the joke" and then determine the time relationships between the panels. The reverse order for this process also occurred, but until the subjects "got the joke" they were never quite confident with their arrangement, and at times would even rearrange a correct order to see if they could discover what was supposed to be funny. Whether or not a test of one's sense of humor is a useful assessment is open to question, but clearly the PSF as developed to date is testing such an "aptitude" at a far lower level than it is capable.

2) Minority Testing

A second broad area of potential for the PSF is within minority testing. Two groups will be considered: the hearing-impaired and ethnic minorities. For the assessment of those with hearing disabilities a nonverbal, visual test is obviously preferrable. Levine (1971) noted that presently there is no adequate intelligence test for the deaf.

Her (1974) assessment of the art indicates that even though the Wechsler performance tests are a heavy favorite, the general opinion of them is one of "muted pessimism", preferred only as "better than most." What is needed is a battery of tests with at least some standardization on a national sample of deaf subjects. Since the current "typical" (or at least relatively common) profile of the deaf child reflects a high Object Assembly but a low PA subtest score (Aimmerman and Woo-Sam, 1972), the potential value of PSF strategies is uncertain. However, the existing research has used the present short, complex picture arrangement subtest. What is proposed instead is a large, multidimensional, group-administered PSF test. Such a test, which is entirely nonverbal and capable of being administered through pantomime, might greatly increase the chances for a contribution to the final battery.

In regard to the testing of ethnic and class minorities, the controversy over bias within intelligence testing has been raging (and will probably continue) for many year (Loehlin, Lindzey, and Spuhler, 1975; Samuda, 1975). The proliferation of nonverbal, pictorial inventories has in fact been related to the need to fairly test those groups which have not developed the particular linguistic styles of the majority, white, middle-class culture. Since Blacks and Whites use different linguistic, speaking styles (Labov, 1970; Stewart, 1970; Baratz, 1973; Hall and Freedle, 1973), in which the same words do not necessarily mean the same thing, (Taylor, 1971), it was largely thought that much test bias would reflect the verbal content and nature of test items. Quay (1971), however, failed to find any effect through the translation of the Stanford-Binet into Black English. In fact, Flaugher (1970) and Mandler & Stein (1971) note that the largest discrepancies are often found on nonverbal items.

Most comparative studies concern Black-White differences and these usually involve Wechslerian tests. In a review of these studies Loehlin, et. al. (1975) found that two-thirds reported Blacks as relatively higher on the verbal than performance items. In a combined ranking of those studies which reported subtest performance, the Block Design received the worst performance, the Picture Arrangement the next worst, and the Picture Completion third worst (with best performance on verbal subtest).

With data such as the above, it may be ironic that the attempts to develop culture-free or culture-fair tests have tended to emphasize performance on nonverbal items, but it is at least understandable why they have generally failed. For example, Cattell and Cattell (1960) developed the Cattell Culture Fair Intelligence Test in 1944. It is a nonverbal test using abstract symbols in the style of Raven's matrices. Reviews by Tannenbaum (1965) in Buros, Milholland (1965), Willard (1968), and Kidd (1962) are reflective of the general opinion that the test was not successful in removing cultural or ethnic performance discrepancies. Kidd (1962) and Marquart and Bailey (1955) went so far as to conclude that a culture-free test is impossible to develop.

Raven's various matrices tests (Raven, 1960; Court, 1972) have often been cited as culture fair tests, again largely because of the absence of explicit, readily observable cultural differences within nonverbal items. However, reviews by Burke (1958) and Irvine (1966) question its "fairness" across cultures. Goodenough's Draw-A-Man (1962) is another example of a "nonverbal-therefore-not-biased" misnomer, yet Goodenough and Harris themselves have stated that it was not free from cultural bias and that the search for such a test is illusory (Anastasi, 1968).

The main problem, of course, is that a test score is influenced by a multitude of factors other than the content. Fishmen, et. al. (1964),

Samuda (1975), and Zigler and Butterfield (1968), listed a number of variable which would differentiate cultures in test performance: motivation, test-taking experience, self-concept, test anxiety, test environment (e.g., ethnicity of tester), speededness of test, relevancy, expectations, etc. Ability tests in large part are inevitably measurements of the learning experience of the individual, and to expect that a test will not reflect the different learning experiences within cultures is a bit naive (Wesman, 1968). Developing a test which can predict useful social criteria but doesn't reflect differential cultural experiences (i.e., isn't heavily weighted by the skills valued in the predominant culture), would be extremely difficult (Bennet, 1970). The conclusion seems to be that no test is culture fair (Anastasi, 1961; Vernon, 1969), nor can one be developed.

An interesting attempt involving a PSF was made by the Chicago group of Davis and Eells in the early fifties. They developed the Davis-Eells Test of General Intelligence or Problem Solving Ability (Davis and Eells, 1953). It used a cartoon format with four types of problems: one in which a social scene was depicted and the subject chose from among three statements (read outloud by the examiner) the one that best applied, another in which three panels depicted alternate means of solving some sort of practical problem (subject chose which represented the best solution), a third concerning pictorially illustrated math problems, and the fourth a pictorial analogies test. The first rwo certainly are relevant as an expansion of the PSF. Although the test was cited by its authors as being culture-fair, this was largely rejected by researchers (Ludlow, 1956; Noll, 1960; Stablein, Willey, and Thomson, 1961), and it has since been taken off the market. However, it should not be entirely

dismissed. Davis and Eells themselves restricted its application to "children coming from the American culture typical of cities and small towns" (1953, p. 1). Scores of rural children and those using a foreign language therefore, were to be interpreted with caution. Their overstatement lay in the claim that it used scenes "equally familiar to all American cultural groups" (p. 8). Minority ethnic group characters and situations are completely absent; there is no representative sample of scenes within the lower socio-economic strata. In view of the many variable preventing the development of a useful culture-fair test, it is not surprising that this effort failed.

We would suggest, however, a consideration of the PSF's suitability within a culture-specific format. In the same sense that the WISC or WAIS is specific to the majority culture, and is useful for identifying exeptional talent within it, so too could a test be specific to a minority culture (Williams, 1972). Two well-known verbal information tests have been developed: Williams' (1972) Black Intelligence Test of Cultural Homogeneity and Adrian Dove's Counterbalance General Intelligence Test developed in 1966. Whether or not these tests predict anything very well is unknown due to the lack of relevant research, but they certainly have been successful in reversing the relative scores of cultural groups. If a PSF included minority culture-specific scenes and events, as well as the majority's, the comparative scores of members of various cultures could be analyzed. There has been some talk concerning the content bias of Wechsler's pictorial tests (Hewitt and Massey, 1969; Breiger, 1956; Matarazzo, 1975) but few, if any, controlled studies.

It is highly doubtful, then, that a PSF with varying culture-specific contents would be culture-fair, but certainly it would be useful for

advancing the knowledge concerning different cultural performances. In the least, the PSF style presents a very appropriate and ready format for representing various cultural groups. The efforts to integrate verbal and mathematical items with different cultural groups is often difficult, artificial, and at times meaningless "tokenism." The use of social scenes in pictorial format offers a test setting with a far greater capacity for comprehensive and meaningful representation.

3) Occupational and Practical Assessment

The attempt in the Davis-Eells test to tap "problem-solving capacity" in real-life situations, leads to the third major area of potentiality, the testing of job success and other practical assessments. Present intelligence tests are not very successful in predicting success within an occupation. Nor are actual accomplishments in social leadership, the arts, science, music, etc., highly correlated with mental intelligence. Thorndike and Hagen (1959) and Ghiselli's (1966) major works document these points. One of the major difficulties in developing items for more practical tests is a need to adequately set up or describe relevant, everyday problems in a manner which depicts their ramifications and complexities but without becoming so lengthy as to be testing basically reading comprehension and verbal facility. The capability of a PSF to present a complex social scene within a single, nonverbal stimulus should make this format a candidate for serious consideration. As stated earlier, Guilford and associates have already experienced some limited success in the area of social knowledge. Joseph King's application of single-picture stimuli in some of his Job-Tests series (1948-1960) is the beginning of a similar application within a more practical realm.

4) A Benign Test Strategy

A final area of potential is the PSF's potential as a benign test strategy, seen by subjects as inherently entertaining and nonthreatening. It has been the experience in informal data collection (i.e., the administration of the full WISC-R and WAIS, and the trial administrations of our own PSF items) that the PSF is considered to be one of the most enjoyable and anxiety-reducing subtests. Its value in this regard may not be so great as to resolve the many motivational, text anxiety, and negativism problems cited for minority testing, but it should serve to combat them (e.g., Strang, Bridgement, and Carrico, 1974). In the Davis-Eells Studies, for example, one of the major criteria for the inclusion of a test was its potential for an intrinsic stimulation other than "academic competitiveness." "Since these mental problems interest children, they are more effective in stimulating the use of their full problem-solving capacity than for those problems which seem unreal.... This test stimulates interest, rather than the boredom or actual fear so frequently aroused by tests" (Davis and Eells, 1953, p. 4). Coupled with a meaningful cultural representation (and variety) it could be quite a powerful device.

It is interesting to note that Davis and Eells considered a pure PA format, but eliminated it after three trial runs. The reason for the rejection was two-fold. First, their test did not appear to possess the realistic, face-validity stimulation that the other tests aroused. In reporting this, they observe "This is hardly confronted with a series of events out of their proper chronological sequence which he

must sort into the proper order" (p. 45). We feel that this finding was a result of their specific test models, rather than to a limitation of PSF generally. Our present recommendations for the further development of the PSF include styles not reflected the Davis and Eells test (in fact, our own preference is for the Missing Cartoon format, due to its relative ease in administration and scoring).

Davis and Eells' second criticism concerned a difficulty in getting the children to understand exactly what they were supposed to do. For younger children the pure PSF was at times conceptually difficult to grasp, especially within a group-administered format. It was for this reason, also, that the PAS was eliminated from the Wechsler Preschool and Primary Scale of Intelligence (Wechsler, 1967, p. 8). There probably is, then, a lower-bound limit on a group PSF. But the PA form is only one among many PSFs, and whatever the lower bound might be it should be empirically established.

Conclusions

On the whole, the PSF would seem to have potential for a wide variety of areas which remain untapped. Those expansions that have occurred have tended to come in the area of personality testing. Craig (1969) and Rapaport, Gil, and Schafer (1968) have both developed separate systems for projective interpretations of the Wechsler PA items, and Tomkins and Horn (Tomkins and Miner, 1957) developed a comprehensive group-administered projective device using a PSF. This potential for making projective interpretations of a subject's response to pictorial social scenes is rather clear and has never ceased to be

exploited (Taylor, 1961; Sattler, 1974), whether in the traditional form of the Thematic Apperception Test or in variations such as the Family Attitudes Test (Jackson, 1966). It is regrettable that there has been no similar expansion of the PSF within intelligence testing.

There are definite limitations and difficulties inherent in developing a PSF. In terms of content specificity, for example, especially if one devises culture-specific items, one runs the risk of having the test become outdated (i.e., lose its appeal) in a relatively short period of time. The depictions of specific roles and situations can quickly become obsolete or sources of embarrassed prejudice in subsequent years (e.g., a test developed ten years ago would probably not be nearly as sensitive to sexist references as it would be today). A PSF also entails construction difficulties not found within other test formats, and this may contribute to its neglect. The scientist-scholar finds much less effort and perhaps more enjoyment in devising verbal, mathematical, and spatial material. The need to think up the detail for a pictorial series, to try to draw them, or to pay the expense of hiring an illustration makes the task inherently unattractive. Finally, little is known of the extent to which cartoons are inherently unfair towards certain groups. One can anticipate such problems as blind subjects, but there are other problems, such as perhaps different cultures have a differential exposure to cartoons (Machover, 1943; Mandler and Stein, 1977).

These concerns, however, do not in our view justify the present lack of further development. The breadth and depth of the potentialities for PSF in terms of the reflections of such content domains as moral

judgment and humor; unique the testing and representation of certain minorities, the testing for practical, social, and occupational knowledge, and the reduction of test-anxiety, motivation, and negativism, are quite exciting. Some of the so-called difficulties of the PSF (e.g., its complexity and vicissitudes) are in fact its assets. It presents a format readily appealing yet open for testing a wide variety of aptitudes in a variety of forms. We urge a reconsideration of its merits and the attention of researchers and test developers.

References

Adkins, D.C. and Lyerly, S.B. Factor Analysis of Reasoning Tests. Chapel Hill, N.C.: University of North Carolina, 1951.

Anastasi, A. "Psychological tests: uses and abuses." Teachers College Record, 1961, 62, 389-393.

Anastasi, A. "Culture fair testing." In N.E. Gronlund (Ed.) Readings in Measurement and Evaluation, N.Y.: Macmillan, 1968, 280-286.

Baratz, J. C. "Language abilities of black Americans." In K.S. Miller and R.M. Dreger (Eds.) Comparative Studies of Blacks and Whites in the United States, N.Y.: Seminar Press, 1973, 125-183.

Bennett, G.K. "Response to Robert Williams." Counseling Psychology, 1970, 2, 2, 88-89.

Blatt, S.J. and Allison, J. "The intelligence test in personality assessment." In A.I. Rabin (Ed.) Projective Techniques in Personality Assessment, N.Y.: Springer, 1968, 421-460.

Blatt, S.J. and Quinlan, P. "Punctual and procrastinating students: a study of temporal parameters." Journal of Consulting Psychology, 1967, 31, 169-174.

Breiger, B. "The use of the Wechsler-Bellevue picture arrangement subtest as a projective technique." Journal of Consulting Psychology, 1956, 20, 132.

Brown, A., Stein, S., and Rohrer, P. Manual of Directions for the Chicago Non-verbal Examination, N.Y.: The Psychological Co., 1936.

Burgemeister, B., Blum, L., and Lorge, I. Columbia Mental Maturity Scale, 3rd Edition, N.Y.: Harcourt Brace Jovanovich, 1972.

Burke, H. "Raven's Progressive Matrices: a review and critical evaluation." Journal of Genetic Psychology, 1958, 93, 199-228.

Cattell, R. and Cattell, A. Manual for the Cattell Culture-Fair Intelligence Test, Indianapolis: Bobbs-Merrill Co., 1960.

Cohen, J. "The factorial structure of the WISC at ages 7-6, 10-6, and 13-6." Journal of Consulting Psychology, 1959, 23, 285-299.

Cornell, E. and Coxe, W. Cornell-Coxe Performance Ability Scale, Examination Manual, Yonkers-on-Hudson, N.Y.: World Book Co., 1934.

Court, J. H. Researchers' Bibliography for Raven's Progressive Matrices and Mill Hill Vocabulary Scales, Adelaide, 1972.

Craig, R. J. "An illustration of the Wechsler Picture Arrangement subtest as a thematic technique." Journal of Projective Techniques and Personality Assessment, 1969, 33, 286-289.

Cross, K. P. Education for Diversity: New Forms for New Functions. Paper presented at the National Conference on Higher Education, Chicago, 1974.

Davis, A. and Eells, K. Davis-Eells Test of General Intelligence or Problem-Solving Ability, Manual. N.Y.: World Book Co., 1953.

Decroly, O. "Epreuve nouvelle pour l'examen mental." L'Annes Psychologique, 1914, 20, 140-159.

Dickstein, L. and Blatt, S. "Death concern, futurity, and anticipation." Journal of Consulting Psychology, 1966, 30, 11-17.

Dickstein, L. and Blatt, S. "The WAIS picture arrangement subtest as a measure of anticipation." Journal of Projective Techniques and Personality Assessment, 1967, 31, 3, 32-38.

Dingman, H. and Myers, C. "The structure of intellect in the mental retardate." In N. R. Ellis (Ed.) International Review of Research in Mental Retardation, Vol. I. N.Y.: Academic Press, 1966, 55-76.

Dunn, L.M. Expanded Manual for the Peabody Picture Vocabulary Test. Minneapolis: American Guidance Service, 1965.

Fishman, et. al. "Guidelines for testing minority group children." Journal of Social Issues Supplement, 1964, 20, 129-145.

Flaugher, R.L. "Testing practices, minority groups, and higher education: a review and discussion of the research." Research Bulletin 70-41, Princeton, N.J.: ETS, 1970.

French, J. Manual: Pictorial Test of Intelligence. Boston, Houghton Mifflin, Co., 1964.

Ghiselli, E. The Validity of Occupational Aptitude Tests. N.Y.: Wiley, 1966.

Goodenough, F.L. Measurement of Intelligence by Drawings. Yonkers-on-Hudson, N.Y.: World Book, 1926.

Graham, E. and Shapiro, E. "Use of the Performance Scale of the Wechsler Intelligence Scale for Children with the deaf child." Journal of Consulting Psychology, 1953, 17, 396-398.

Guilford, J.P. Personality, N.Y.: McGraw-Hill, 1959.

Guthrie, G., Tayag, A., and Jacobs, P. The Philippine Nonverbal Intelligence Test (Rb No. 70). University Park, Pennsylvania: Department of Psychology, Penn. State University, Feb., 1969.

Hall, W. and Freedle, R. "A developmental investigation of standard and nonstandard English among black and white children." Human Development, 1973, 16, 440-464.

Heussestamm, F. and Hoepfner, R. Inter-Perception Test, Manual. Monitor, 1969.

Hewitt, P. and Massey, J. Clinical Clues from the WISC. Palo Alto, Cal.: Consulting Psychologists Press, 1969.

Hoepfner, R. "The validity of tests of social intelligence." CSE Report No. 85, April, 1973.

Hoepfner, P. and O'Sullivan, S. "Social intelligence and IQ." Educational and Psychological Measurement, 1968, 28, 339-3

Irvine, S. "Towards a rationale for testing attainments and abilities in Africa." British Journal of Educational Psychology, 1966, 36, 24-32.

Jackson, L. Family Attitudes Test. Brussels: Editest, 1966.

Joslin, H. "Intelligence and social awareness: an investigation of the relationship between social awareness, planfulness, and performance on the Wechsler picture arrangement subtests." Dissertation Abstracts International, 1966, 2871-B.

Kagan, J., Rosman, B., Day, D., Albert, J., and Phillips, W. "Information processing in the child" significance of analytic and reflective attitudes." Psychological Monographs, 1964, 78 (1, Whole No. 578).

Kaufman, A. "Factor analysis of the WISC-R at 11 age levels between 6-1/2 and 16-1/2 years." Journal of Consulting and Clinical Psychology, 1975, 43, 2, 135-147.

Kidd, A. "The culture-fair aspects of Cattell's test of g: culture-free." Journal of Genetic Psychology, 1962, 101, 343-362.

King, J. Job-Tests Program--Factor Aptitude Series. Industrial Psychology, Inc., 1948-1960.

Klinger, D. and Saunders, D. "A factor analysis of the items for nine subtests of the WAIS." Multivariate Behavioral Research, 1975, 10, 131-154.

Krippner, S. "WISC comprehension and picture arrangement subtests as measures of social competence." Journal of Clinical Psychology. 1964, 20, 366-367.

Labov, W. "The logic of nonstandard English." In F. Williams (Ed.) Language and Poverty. Chicago: Markham Publishing Co., 1970. 1970, 153-189.

Levine, E. "Mental assessment of the deaf child." The Volta Review, 1971, Feb., 80-105.

Levine, E. "Psychological tests and practices with the deaf: a survey of the state of the art." The Volta Review, 1974, May, 298-319.

Loehlin, J., Lindzey, G., and Spuhler, J. Race Differences in Intelligence. San Francisco: W. H. Freeman and Co., 1975.

Ludlow, H. G. "Some recent research on the Davis-Eells games." School and Society, 1956, 84, 146-148.

Machover, S. "Cultural and racial variations in patterns of intellect." Teachers College, Columbia University Contributions to Education, No. 875, 1943.

Mandler, J. and Stein, N. "The myth of perceptual defect" sources and evidence." Psychological Bulletin, 1977, 84, 1, 173-192.

Marquart, D. and Bailey, L. "An evaluation of the culture-free test of intelligence." Journal of Genetic Psychology, 1955, 86, 353-358.

Matarazzo, J. Wechsler's Measurement and Appraisal of Adult Intelligence, 5th and Enlarged Edition, Baltimore: William and Wilkins Co., 1972.

Matarazzo, J. and Wiens, A. "Black Intelligence Test of Cultural Homogeneity and WAIS scores of black and white police applicants." Journal of Applied Psychology, 1977, 62, 1, 57-63.

McCay, V. and Brown, D. "A guide to psychological tests and testing procedures in the evaluation of deaf and hard-of-hearing children." Journal of Speech and Hearing Disorders, 1964, 29, 4, 414-423.

McDonald, K. G. and Crookes, T.G. "The WAIS picture arrangement test in British psychiatric patients." British Journal of Social and Clinical Psychology, 1967, 6, 72.

McMurray, R. and Arnold, P. S.R.A. Pictorial Reasoning Test, Examiner's Manual. Chicago: Science Research Associates, 1973.

Meeker, M. The Structure of Intellect. Columbus, Ohio: Charles E. Merrill, 1969.

Milholland, J. "Culture-fair Intelligence Test." In Oscar Buros (Ed.) Sixth Mental Measurement Yearbook. Highland Park, N.J.: Gryphon Press, 1965, 719-721.

Myklebust. Development and Disorders of Written Language, Vol. I, Picture Story Language Test. N.Y.: Grune and Stratton, 1965.

Noll, V. H. "Relation of scores on Davis-Eells games to socio-economic status, intelligence test results, and school achievement." Educational and Psychological Measurement, 1960, 20, 119-129.

Orpet, R. and Meyers, C. "Discriminant function analysis of conservation stages by structure of intellect and conceptual style variables." Proceedings of the 78th Annual Convention of the American Psychological Association, 1970, 5, 279-280.

O'Sullivan, N. and Guilford, J. P. "Six factors of behavioral cognition: understanding other people." Journal of Educational Measurement, 1975, 12, 4, 255-271.

Otis, A. Otis Group Intelligence Scale. Yonkers-on-Hudson, N.Y.: World Book Co., 1929.

Quay, L. "Language dialect, reinforcement, and the intelligence-test performance of Negro children." Child Development, 1971, 42, 5-15.

Rapaport, D., Gil, M., and Schafer, R. Diagnostic Psychological Testing (Rev. Ed.) N.Y.: International Universities Press, 1968.

Raven, J. Guide to the Standard Progressive Matrices, Sets A, B, C, D, and E. London: H.K. Lewis and Co., Ltd., 1960.

Rulon, P. "A semantic test of intelligence." (1952). In Anastasi (Ed.) Testing Problems in Perspective. Washington, D.C.: American Council on Education, 1966, 472-480.

Samuda, R. Psychological Testing of American Minorities: Issues and Consequences. N.Y.: Dodd, Mead, and Co., 1975.

Sattler, J. Assessment of Children's Intelligence. Philadelphia: W. B. Saunders Co., 1974.

Schill, T. "The effects of MMPI social introversion on WAIS picture arrangement performance." Journal of Clinical Psychology, 1966, 22, 1, 72-74.

Schill, T., Kahn, M., and Muehleman, T. "Verbal conditionability and Wechsler picture arrangement scores." Journal of Consulting and Clinical Psychology, 1968a, 32, 6, 718-721.

Schill, T., Kahn, M., and Muehleman, T. "Wechsler Adult Intelligence Scale picture arrangement performance and participation in extra-curricular activities." Journal of Clinical Psychology, 1968b, 24, 1, 95-96.

Shanley, L.A., Walker, P., and Foley, J. "Social intelligence: a concept in search of data." Psychological Reports, 1971, 29, 1123-1132.

Sigel, I. "How intelligence tests limit understanding of intelligence." Merrill-Palmer Quarterly, 1963, 9, 39-56.

Solway, K. S., Cook, T., and Hays, J.R. "WISC subtest patterns of delinquent female retardates." Psychological Reports, 1976, 38, 42.

Solway, K.S., Fruge, E., Hays, J., Cody, J.A., and Gryll, S. "A comparison of the WISC and WISC-R in a juvenile delinquent population." The Journal of Psychology, 1976, 94, 101-106.

Solway, K.S., Hays, J., Roberts, T.K., and Cody, J. "Comparison of WISC profiles of alleged juvenile delinquents living at hom versus those incarcerated." Psychological Reports, 1975, 37, 403-407.

Stablein, J., Willey, D., and Thomson, C. "An evaluation of the Davis-Eells (Culture-Fair) Test using Spanish and Anglo-American children." Journal of Educational Sociology, 1961, 35, 73-78.

Stewart, W. "Towards a history of American Negro dialect." In F. Williams (Ed.) Language and Poverty. Chicago: Markham Press, 1970.

Strang, H., Bridgeman, B., and Carrico, M. "Effects of 'game' versus 'test' task defintion for third-grade children on three subtests of the WISC." Journal of Educational Measurement, 1974, 11, 2, 125-128.

Tannenbaum, A. "Culture-fair Intelligence Test." In Buros (Ed.) Sixth Mental Measurements Yearbook. Highland Park, N.J.: Gryphon Press, 1965, 721-723.

Taylor, E. M. Psychological Appraisal of Children with Cerebral Defects. Cambridge, Mass.: Harvard University Press, 1961.

Taylor, O.L. "Some sociolinguistic concepts of black language." Today's Speech, 1971, 19-26.

Tenopyr, M.L. "Social intelligence and academic success." Educational and Psychological Measurement, 1967, 27, 961-965.

Thorndike, R. L. and Hagan, E. 10,000 Careers, N.Y.: Wiley, 1959.

Tomkins, S. and Miner, J. The Tomkins-Horn Picture Arrangement Test. N.Y.: Springer Publishing Inc., 1957.

Turner, R., Willerman, L., and Horn, J. "Personality correlates of WAIS performance." Journal of Clinical Psychology, 1976a, 32, 2, 349-354.

Turner, R. and Horn, J. "MMPI item correlates of WAIS subtest performance." Journal of Clinical Psychology, 1976b, 32, 3, 583-594.

Turner, R., Willerman, L., and Horn, J. "A test of some predictions from the personality assessment system." Journal of Clinical Psychology, 1976c, 32, 3, 631-643.

Vermeylen, P.E. "Ability factors and environment influences." American Psychologist, 1965, 20, 723-733.

Wechsler, D. The Measurement and Appraisal of Adult Intelligence (4th edition). Baltimore: Williams and Williams, 1958.

Wechsler, D. Wechsler Preschool and Primary Scale of Intelligence, Manual. N.Y.: The Psychological Corporation, 1967.

Wesman, A. G. "Intelligence testing." American Psychologist, 1968, 23, 267-274.

Willard, L.S. "A comparison of Culture-Fair Test scores with group and individual intelligence test scores of disadvantaged Negro children." Journal of Learning Disabilities, 1968, 1, 584-589.

Winne, J. and Gittinger, J. "An introduction to the Personality Assessment System." Journal of Community Psychology, Special Monograph Supplement, 1973, 1, 99-163.

Yerkes, R. (Ed.) Memoirs of the National Academy of Sciences. Washington: Government Printing Office, 1921.

Yoakum, C. and Yerkes, R. (Eds.) Mental Tests in the American Army. London: Sidgwick and Jackson, Ltd., 1920.

Zigler, E. and Butterfield, E. C. "Motivational aspects of changes in IQ test performance of culturally deprived school children." Child Development, 1968, 39, 1-14.

Zimmerman, I.E. and Woo-Sam, J. "Research with the Wechsler Intelligence Scale for Children: 1960-1970." Psychology in the Schools, Special Monograph Supplement, 1972, 9, 232-271.

APPENDIX

(Guru Cartoon)

see Kovarsky, Anatol. Kovarsky's World. New York: Knopf 1956.

Appendix was removed from this document prior to its being submitted to the ERIC Document Reporduction Service.